

## SYSTEM FOR PARENTAL CONTROL IN VIDEO PROGRAMS BASED ON MULTIMEDIA CONTENT INFORMATION

### BACKGROUND OF THE INVENTION

#### Field of the Invention

The current invention relates to parental control systems for entertainment systems. More specifically, the present invention relates to parental control systems for dynamically filtering the content of a video program in real time on a segment by segment basis, based on at least one user-supplied category of filtering criteria. Methods for dynamically filtering the content of a video program in real time on a segment by segment basis are also disclosed.

#### Discussion of Related Art

In current television and video programs the rating of such programs is based on subjective program content evaluation generated by human observation, i.e., based on some sort of interpreted standard for classification. The result of this content evaluation is a set of recommendations, such as ratings. Examples include ratings of G, PG, PG-13, R, NC17, X for movies, or TV-Y, TV-Y7, TV-G, TV-PG, TV-14, and TV-MA for television programs. Such ratings are typically shown at the beginning of a video program, perhaps with an additional set of sub-classifications, such as "V" for violence, "AL" for adult language, and "N" for nudity, for example.

One of the limitations of this process, i.e., subjective evaluation, is that it does not provide a comprehensive set of criteria for parents to judge the appropriateness, or lack thereof, of TV programs. For example, the Motion Picture Association of America (MPAA) describes a PG rating for movies as referring to a film which a parent should view or at least get more detailed information on before letting their children watch the movie. The MPAA further states that PG

movies may contain some profanity, some violence, and/or some brief nudity – in some combination possibly, but no drug use. Such a description is very vague and uncertain, and does not provide enough detail for parents to make an educated decision about whether some or all of a PG-rated program may be appropriate for their children. For example, while a given set of parents may not find brief nudity involving buttocks objectionable, that same group of parents may be very adverse to having their children view bare torsos, certain types of violence, or hear particular forms of profanity. Another set of parents may have a completely different set of criteria to determine whether a PG rated movie is acceptable for their children to watch. Further, standard rating systems do not address other potentially objectionable material such as social situations (e.g. discussions of homosexuality), philosophy (e.g. pro-life or pro-choice) or matters of religion.

Furthermore, traditional rating systems only apply to video programming which has been evaluated by some entity empowered to do so. Ratings are not assigned to home videos, some limited distribution video programs sold via mail order, etc.

Current parental control techniques regarding video programming rely either on overall program ratings or the matching of program identification against a database of approved or restricted material. For example, televisions currently being produced include technology known as the “V-Chip”, which uses rating information inserted in the broadcast television video programs to determine, based on user supplied information, whether the program is viewable or not. Such rating information is found in line 21 of the vertical blanking interval (VBI) of each field of a frame of the video program. It will be noted that line 21 of the VBI contains closed caption information for standard analog television broadcasts, possibly even in multiple languages, and also extended data services (XDS) packets, i.e., the rating information for a program is stored among other data. This mechanism for transporting data in line 21 of the VBI is described in the ANSI EIA-608 specification, and a summary of VBI line 21 data interpretation can be found in the January

2000 issue of Nuts & Volts. A different specification discusses data encoding for digital video playback and broadcasts.

U.S. Patent No. 4,930,158 to Vogel discloses a selective video playing system in which a video program on a video tape contains an embedded classification code. That code can then be used to inhibit replay of the video program, under user control. U.S. Patent No. 4,930,160 to Vogel discloses automatic censorship of video programs, and also involves the use of stored classification codes – either a single code or a group of codes, but with the program coming from either video tape or via broadcast. When a restricted classification code is detected, an alternative video source is automatically switched in. With these devices and methods, as well as the V-chip solution, parents are limited to employing predetermined classifications, which, as previously indicated, may not agree with their particular values.

An alternative approach to parental control has been to try in some fashion to identify a given video program and based on that identification, make a determination whether or not the program should be blocked or be permitted to be viewed. Several techniques for this identification have previously been proposed.

U.S. Patent No. 4,639,779 to Greenberg discusses a method by which a unique ID is encoded into a video program to allow the program to be automatically identified, while U.S. Patent No. 4,843,562 to Kenyon identifies video programs by storing a spectrogram “fingerprint” of the program into a library, and then using that library to look up future programs via their spectrogram to try and find a match. U.S. Patent No. 5,019,899 to Boles uses a generated digital signature for video information to try and match it to a database.

In contrast, U.S. Patent No. 5,485,518 combines a video program identification technique with program blocking, by enabling a parent or other user to prevent someone from watching a

program which is not in an "approved program" database. So, if the program is not in the database, or if it is in the database but meets certain blocking criteria, its complete viewing would be blocked. The video program identification technique relies on video and audio signal processing.

5

It should be noted that none of the existing methods implemented in parental control systems permits selective, automatic filtering of only the offending portions out of a video program, but instead merely provides for blocking of entire programs either based on generic, non-specific ratings, or by identifying the program and making a determination whether such program may be viewed.

10

What is needed is an automatic system capable of evaluating video programs which filters, blocks, or masks portions of the video programs according to user-supplied criteria, in real or substantially real time, i.e., as the video program is playing or being broadcast. What is also needed is a parental control system, which learns the user-supplied criteria. It would be desirable if the parental control system could, for example, filter video programs using criteria appropriate for different groups of viewers.

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179

## SUMMARY OF THE INVENTION

Based on the above and foregoing, it can be appreciated that there presently exists a need in the art for a parental control system which overcomes the above-described deficiencies. The present invention was motivated by a desire to overcome the drawbacks and shortcomings of the presently available technology, and thereby fulfill this need in the art.

The present invention provides an automatic system to evaluate video programs using criteria supplied by a user (a parent, for example) via the multimedia content, e.g. visual, audio, and textual content, of video programs and which then filters, blocks, or masks portions of the video programs according to those criteria, in real time. Such blocking or masking may include simply skipping over the material determined to meet the user specified criteria, substituting an alternate "safe" signal for the duration of the offending program segment, or masking portions of the video or audio content, e.g. blurring a naked body or garbling profanity.

Preferably, the parental control system according to the present invention includes one or more multimedia processors, which analyze incoming visual, audio, and textual, content and compare the resultant analysis to specific user-specified or selected criteria. For example, in the visual content one can have expressions (e.g. facial and body), behaviors (e.g. shooting a gun, sexual activity, driving a vehicle), body attributes (e.g. skin tone or shape), violence (e.g. explosions, fighting), etc. In the audio domain one can have sound level (e.g. heavy sound with a lot of low frequency noise for an explosion), verbal expressions (e.g. profanity, slurs, slang, cursing, innuendo), "heavy breathing" sounds (e.g. as might occur during a sex scene), etc. In the textual domain there is semantic content which can include adult language, themes, etc. Raw textual information applied to the textual domain can be extracted from the program in a number of methods. One such method would be the use of speech recognition technology applied to the audio information stream from the program. For television programming, another method to extract the

raw textual information would be to use closed caption information. Both methods advantageously could be employed in the same parental control system, either to complement each other and/or provide redundancy.

5 In one aspect, the present invention provides a method for dynamically filtering the content of a multimedia program in real time on a segment- by- segment basis responsive to a filter criteria, comprising extracting audio, video, and transcript features from segments including the multimedia program, generating a numeric ranking for each filter criteria for each applicable filter category (e.g. violence, nudity, religion, etc.) each of the segments, and when  
10 the combined respective numeric rankings for that segment exceeds a threshold, processing that segment to thereby eliminate material corresponding to the filter criteria. Preferably, the filter criteria corresponds to language included in the segment being processed, and the audio portion of the segment is modified during the processing step. Alternatively, the filter criteria corresponds to an image included in the segment being processed, and the video portion of the segment is either modified or skipped during the processing step. In an exemplary embodiment, the numeric ranking is a weighted numeric ranking. In that case, each weighting factor employed in generating the weighted numeric ranking identifies a characteristic such as age or religious affiliation of a respective viewer of the multimedia program. In any event, the numeric ranking for each applicable filter category in each segment is generated by comparing the content  
15 of each segment to the filter criteria.  
20

It should be mentioned that, in appropriate cases, i.e., when the combined numeric rankings for proximate ones of the segments each exceed the threshold, the method preferably merges the proximate ones of the segments and any intervening segments to thereby produce a  
25 merged segment. In that situation, the processing step permits processing the merged segment to thereby eliminate material corresponding to the filter criteria. Moreover, the filter criteria preferably includes first and second filter criteria, the generating step includes generating first

and second numeric rankings for respective first and second filter criteria for each of the segments, the method including the further steps of when the respective first numeric ranking for that segment exceeds a first threshold, processing that segment to thereby eliminate material corresponding to the first filter criteria, when the respective second numeric ranking for that segment exceeds a second threshold, processing that segment to thereby eliminate material corresponding to the second filter criteria .

In another aspect, the present invention provides a parental control system filtering objectionable material from a multimedia program in accordance with a filter criteria. Preferably, the parental control system includes a transcript analysis module extracting first audible features and text from a sequence of segments included in the multimedia program, a visual analysis module extracting video features from the sequence of segments included in the multimedia program, an audio analysis module extracting second audible features from the sequence of segments included in the multimedia program, an analyzer which generates a combined numeric ranking for each of the segments and which generates a respective control signal when the combined numeric ranking exceeds a threshold, and a filter which processes one of the segments of the multimedia program in response to a received respective control signal. The filter either modifies one of the first and second audible features of the respective segment, or modifies the video feature of the respective segment, or eliminates the respective segment from the filtered multimedia program output by the parental control system. Preferably, the parental control system includes a learning module. In that case, selected ones of the first audible features and text extracted by the transcript analysis module, the video features extracted by the visual analysis module, the second audible features extracted by the audio analysis module and user data provided by a controlling user of the parental control system are employed by the learning module to generate the filter criteria. In an exemplar case, the learning module includes a neural network. The parental control system advantageously can be incorporated into a television set, a settop box, or a personal video recorder

## BRIEF DESCRIPTION OF THE DRAWINGS

5        These and various other features and aspects of the present invention will be readily understood with reference to the following detailed description taken in conjunction with the accompanying drawings in which:

      Fig. 1 is a high-level block diagram of a parental control system according to the present invention;

10       Fig. 2 is a high-level block diagram of the parental control system according to the present invention illustrated in Fig. 1 depicting operation in a learning mode; and

      Fig. 3 is a block diagram of the parental control system according to the present invention illustrated in Fig. 1 depicting operation in a viewing mode.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

15       In the parental control system and corresponding method according to the present invention, the parental control system, which can be embodied in a television, in a set-top box, or some other type of peripheral, e.g., a personal video recorder (PVR), operates in two related modes of operation. The first mode is a learning mode in which the controlling user (typically a parent or guardian) interacts with the system to configure with examples of the types of video program scenes or segments the controlling user finds objectionable for a selected group of viewers (e.g. small children could be one group, young teenagers another, and grandparents yet another group). In fact, 20       the user might use a selection of extreme examples to train the system coarsely, and then use more discreet examples to fine-tune the training of the system in the learning mode. The second mode of operation of the parental control system and corresponding method is the operational view mode,



where the system is filtering the desired video program according to the criteria it has learned about in the learning mode (and/or which shipped with the system or was generically configured with).

Fig. 1 is a high-level block diagram of a preferred embodiment parental control system according to the present invention, which includes feature extraction circuitry 40, analyzer circuitry 80 and multimedia processor 90. The feature extraction circuitry 40, which will be discussed in greater detail below, receives multimedia content from either tuner 30 or multimedia memory 20. It will be appreciated that memory 20 is preferably connected to analyzer circuitry 80 and multimedia processor 90, which permits the latter devices to employ multimedia memory 20 as buffer memory. The size of the memory 20 is exaggerated to indicate that the memory 20 could be, in other exemplary embodiments, a memory having sufficient capacity to store many hours of multimedia content. It will be readily understood that having a large memory associated with the parental control system would permit the parental control system to process and filter a program, e.g., a feature length film, so as to obtain a film suitable for presentation to a child. Stated another way, a parental control system equipped or connected to a large multimedia memory would permit the controlling user to filter the original film in a myriad of ways to produce versions of the film suitable for several disparate groups in the user's household.

Still referring to Fig. 1, a central processing unit 10 coupled to a memory 15, which advantageously can include both volatile and non-volatile memory, controls each of the multimedia memory 20, tuner 30, feature extraction circuitry 40, analyzer circuitry 80 and multimedia processor 90. As illustrated in Fig. 1, memory 15 can provide memory storage functions for feature extraction circuitry 40, analyzer circuitry 80 and multimedia processor 90. It should be noted that feature extraction circuitry 40 and analyzer circuitry 80 may themselves be or include separate multimedia processors, and/or dispatch processing functions through multimedia processor 90. Memory 15 advantageously may store software suitable for converting a general purpose device such as a personal video recorder (PVR) equipped with a disk drive and several digital signal processors

(DSPs) into a parental control system according to the present invention.

As discussed above, the first mode of operation of the parental control system is a learning mode in which the user interacts with the parental control system to instruct by example, i.e., by providing the parental control system with examples of the types of video program scenes or segments that the controlling user finds objectionable for a selected group of viewers. In an exemplary case, the user might use a selection of extreme examples to train the system coarsely, and then use more discreet examples to fine-tune the training of the system in the learning mode.

Such examples provided by the controlling user (hereafter referred to simply as the "user") would be the realization of filtering criteria established by the user. The parental control system according to the present invention advantageously may be provided with a preprogrammed set of stock segments illustrating a range of criteria – resulting in a "Smart Box" mode of operation. Such stock criteria could be determined regionally (via survey), or according to the ethnic or the religious persuasion of a purchaser, or via any one of a number of different demographics. In one variation on the parental control system and corresponding method, a user could even download a number of categories of criteria into the parental control system from the Internet, either to bypass the user needing to custom configure the parental control system, or in order to provide a better base from which to start custom configuration of the system.

It should be mentioned that the user advantageously can continuously provide the system with input during the learning phase. For example, the user can mark at least one of the audio, visual or textual parts as being objectionable. Preferably, the user provides start and end times, e.g., by marking these times on a timeline showing the video or a timeline enhanced with the keyframes (thumbnail images or the like), or text, or playback of the audio tracks, or combinations thereof. This input can be related to the objectionable categories. In addition, the user can also provide the degree of sensitivity (objectionability), e.g., the section of the timeline containing mildly

objectionable material could be identified by single brackets while truly objectionable material could be identified by double brackets.

The parental control system and corresponding method may include at least two existing, but unconfigured categories: violence; and sexual content. The user can create additional categories for anything the user is concerned about others viewing. Such additional categories could include ones based on religious beliefs, personal philosophies, and political affiliation, among others. All categories need to be populated with criteria to help the parental control system and corresponding methods learn the type of things that are objectionable to the controlling user and what category the objectionable material falls into.

In another variation of the parental control system and corresponding method, segment marking and filtering information can come prepackaged as meta data along with the video program, perhaps provided as part of the XDS component of line 21 of the VBI, or via a parallel data feed. The prepackaged segment marking information would apply to standard categories, or could create a category dynamically. The prepackaged segment marking meta data could be standard for the particular program, or could be custom created on the fly by the program provider based on known user preferences and needs. Another method for the detection of objectionable content is by using a palette of templates that are part of a system database and that highlight common objectionable situations. A template could be a video clip with shooting, an intimate scene, a lady with bare chest, etc. These templates are matched to the input video. This could be essential part of the system to help in the "bootstrapping" i.e. initial phase of using the system. In other words the system would have "stereotypical" objectionable scenes that the user can then mark and grade.

As mentioned above, the second mode of operation of the parental control system and corresponding method is the operational view mode, where the system filters the desired video program according to the criteria that the parental control system has learned during the learning

mode of operation (and/or which shipped with the system or which the system was generically configured with).

In the parental control system and corresponding method, which is illustrated in Fig. 1 and showing in greater detail in Fig. 2, the system consists of three major modules: a feature extraction module, a learning module, and a classification module. The feature extraction and classification modules operate in both the learning mode as well as in the viewing mode, while the learning module is only active during the learning mode of operation. In some exemplary embodiments, however, the learning module may be converted by suitable programming (internal or external) into an analyzer (segmentation and categorization) module, which is discussed below.

Referring to Fig. 2, which depicts the parental control system and corresponding method's operating in the learning mode, the parental control system receives a regular video input 100 stream, which can come from a DVD player, VCR, cable TV feed, satellite, etc., at the input port of video tuner 130 or alternatively video streaming from the Internet. Alternatively, the parental control system can be incorporated into an Internet browser or an e-mail system, i.e., a client software device, since e-mail often includes a Uniform Resource Locator (URL) pointing to a Web site containing objectionable material. Optionally, if the parental control system and corresponding method is enabled to receive program information, as might come from an Electronic Program Guide (EPG), which program information would likely include detailed rating information, the tuner 130 could keep a current copy of the EPG program list 114 on hand, correlating which broadcast program 100 is being viewed via a broadcast channel. Such information would complement or even supplement any rating information, which might be available via the XDS information located on line 21 of the VBI, as used by V-chip technology.

Also integrated with the tuner 130 is input from the user interface 120 used for marking segments by the user. In the parental control system and corresponding method, the user provides

input 124 into the user interface 120 via a remote control (not shown) which has buttons to allow for category and viewer group selection, initiating an objection to a segment of the video input, terminating an objection, freezing a frame, and, using cursor keys or a trackball-type input mechanism, selecting the parts of a visual scene which the user finds objectionable, such as a naked  
 5 body or a gun. It will be appreciated that the same type of mechanism advantageously could be used to highlight objectionable words or phrases in the close captioning information displayed as part of the video input, or in a text output shown by the voice recognition component located in the transcript analysis component 150. For textual input, character strings such as (???), where "???" corresponds to one of the predefined or user-defined categories, might be used in labeling  
 10 categories, the character strings being entered via an on-screen keyboard such as that displayed with a PVR, e.g., via a TiVo recorder, or a keyboard built into the remote control unit.

Preferably, all marking information, whether for a whole segment, individual frame, or audio selection, is transmitted to the tuner 130, where it, along with the video input information and any optional EPG-related information is carried via the transport stream 134 to the demultiplexer 140.

It is the role of the demultiplexer 140 to split the information from the transport stream into the appropriate parts for the three multimedia components of the feature extraction module, represented by the transcript analysis component 150, the visual analysis component 160, and the  
 20 audio analysis component 170. Segment marking information relevant to the multimedia component type is passed along to the respective components by the demultiplexer 140. It will be apparent to those of ordinary skill in the art that the demultiplexer 140 may also include a demodulator to split an NTSC or PAL or similar broadcast signal into respective visual and audible information. Further  
 25 in such case, the demultiplexer 140 would also utilize a frame grabber so that full digitized frames of information can be sent to the visual analysis component 160 of the feature extraction module. If the video input 100 is digital, however, the demultiplexer 140 may have an MPEG-2 or similar

decoder in it, and then would pass the digital visual image data directly to the visual analysis component 160 without needing to use a so-called frame grabber or similar component to digitize the video.

5 In the feature extraction module 40 (composed of the transcript analysis 150, visual analysis 160, and audio analysis 170 components), the audio, visual and transcript parts of the multimedia input are analyzed to "learn" to recognize the objectionable or sensitive segments based on the criteria specified by existing and new user input. Such segments could be parts of news or documentary programs, or action or romantic movies, for example. The sensitive topics normally  
10 revolve around violence and sex but, as previously discussed, the user can introduce new categories. The user will label certain segments with appropriate category labels. For example, the system will be able to learn that fast motion scenes with audio effects (special effects for explosions, sound of fighting and hitting) are generally associated with violent scenes. Nude color from visual and moans and groans in the audio domain are normally associated with sex scenes. There are papers in the  
15 literature dealing with identifying and/or distinguishing naked body parts in an image. See, for example, D. Forsyth and M. Fleck, "Body Plans" (*Proc. IEEE Conf. on Comput. Vis. and Patt. Recog.*, 1997). Moreover, the user can create a new category, e.g., a lifestyle category, so that the user can instruct the parental control system to learn features of multimedia programs that are deemed unsuitable for some or all members of the user's household.

20 The transcript analysis component 150 of the feature extraction module preferably is provided with both the audio stream and any close captioning information, along with any segment marking information specified by the user. The audio stream is translated into text via a speech to text speech recognition subsystem, as is known in the art. The output of the speech recognition  
25 subsystem can be correlated to the close captioning information to determine if any blatant errors in either exist, possibly requiring greater analysis and evaluation of any language elements in the audio stream. Moreover, when a second audio program (SAP) or bilingual close captioning is available

in/for the multimedia program, additional correlations can be made to resolve ambiguities in the transcript analysis component 150 in the feature extraction module 40.

The visual analysis component 160 of the feature extraction module 40 is provided with the digital visual data stream from the demultiplexer 140. The visual analysis component 160 evaluates the incoming digital visual data stream for a broad range of characteristics, including color, texture, objects and object shape, object motion, and scenes.

Low level feature extraction in the video domain includes color (using histograms to map color ranges and popularities), overall motion, and edges. Mid-level features derived from these low-level features include body color and shapes, as well as the visual effects resulting from explosions, objects disintegrating, and gunfire. Ultimately, the features extracted in this component are used to determine whether or not sensitive content exists. For example, color and shape can be used for detection of nudity – different skin tones (for different genotypes) can be detected quite easily in certain color spaces (e.g. HSV) and as such they can be quite meaningful.

As further reference for the technique to implement video data feature analysis, segmentation – and more specifically scene detection - in video data is described in detail in commonly-assigned U.S. Patent Nos., 6,100,941, 6,137,544, and 6,185,363 B1, of which co-inventor of the current invention, Nevenka Dimitrova, is a joint inventor. Object detection and object motion detection in video data is described extensively in U.S. Patent No. 5,854,856, of which the other co-inventor of the present invention, Radu S. Jasinski, is a joint inventor. All of these patents are incorporated herein in their entirety by reference. In addition, description of motion information is also part of the MPEG-7 standard. See, for example, S. Jeannin, R. Jasinski, A. She, T. Naveen, B. Mory, and A. Tabatabai, “Motion Descriptors For Content-Based Video Representation,” (*Signal Processing: Image Communication*, vol. 16, pp. 59-85, 2000).

The audio analysis component 170 of the feature extraction module receives the audio stream from the demultiplexer 140, just like the transcript analysis component 150 does, but processes the stream differently. Low level feature extraction in the audio domain include sound level analysis, LPC, pitch, bandwidth, energy, MFCC (mel cepstral coefficients – used in speech recognition), and Fourier coefficients. Mid-level features derived from low level audio domain features include explosions, objects hitting, object disintegration, heavy breathing, groaning, smacking sounds, and gunfire.

It is important to mention that the “sensitive” segments given to the feature extraction module 40 (i.e. positive examples for the system) should be marked by the user and that the category labels given to these segments are also given by the user, as previously indicated. More specifically, if the user marks a scene from the movie “Terminator” as “violent” then the system will extract all the features from that scene and feed the learning module 180 with each output feature labeled as “violent”. Similarly, nude color objects extracted from the visual domain and moans and groans extracted from the audio domain may be labeled as “sexual”.

In any event, all features extracted by the three feature extraction components, combined with the current criteria specified for the current video signal (e.g., violent, sexual behavior, etc.) are provided to the learning module 180. The learning module 180 preferably employs standard and well-understood learning paradigms to achieve the proper correlation between the labeling of the video input scene and the extracted features. Standard learning paradigms such as Hidden Markov Models (HMMs), Bayesian networks, genetic algorithms, and neural networks advantageously can be employed in the parental control system and corresponding method, combined with nearest neighbor classification, although one of ordinary skill in the art will appreciate that other learning models, or combinations thereof, may be used as well.

The classification module 190 categorizes segments based on whether they are sensitive or



non-sensitive based on the learned categories and the output of the learning module 180. It should be noted that in the parental control system and corresponding method, the user may also want to improve the learning process of the system via the user interface 120 and, therefore, reviews the results of the filtering as output from the classification module 190, and then provides corrections or modifies certain markings, all via the learning module feedback loop 194, back to the learning module 180.

The results of the learning module 180 and classification module 190 are stored in local memory (memory 15 in Fig. 1) in the parental control system. In the parental control system and corresponding method, this would be non-volatile memory.

Fig. 3 shows the viewing mode of operation of the parental control system and corresponding method, which has many functions that are analogous to those in found in the parental control system when it is in the learning mode of operation. Video input 200 represents the video feed into the system from another source, e.g. satellite, cable, DVD, VCR, and is fed into tuner 230, and then, along with any other data the tuner 230 may wish to deliver, is passed on via transport stream 234 to demultiplexer 240. It will be appreciated that the multimedia program can also be output from the multimedia memory 20 illustrated in Fig. 1; the output of multimedia memory 20 being the information carried by transport stream 234. From there, in turn, the video input is parceled out to the feature extraction module 40" components, namely the transcript engine 250, the visual engine 260, and the audio engine 270, where the incoming data is analyzed and processed, much in the same way as the same named components in the learning mode of operation function. In fact, the tuner 230, the demultiplexer 240, the transcript engine component 250, the visual engine component 260, and the audio engine component 270 advantageously can be exactly the same processes and system components as the tuner 130, the demultiplexer 140, the transcript engine component 150, the visual engine component 160, and the audio engine component 170, with the only difference in their operation being that user supplied criteria are not being passed

along through these components in view (second) mode of operation.

Where some noticeable difference occurs is with the segmentation and categorization module 280, which uses the feature extraction from the three feature extraction components, the transcript engine 250, the visual engine 260, and the audio engine 270, in combination with the previously stored learned criteria, in order to determine whether to tell the filtering module 290 whether to filter the video program during a given segment. In an exemplary embodiment, the learning module 170 advantageously can be converted into the segmentation and categorization module 280, and vice versa, by applying suitable software controls to a general-purpose analyzer.

In any event, the segmentation and categorization module 280 preferably determines the beginning and ending time of sensitive scenes, classifies sensitive scenes and, if necessary, merges scenes that are located proximate to one another. For the latter, this can be accomplished by extending the video signal from the video input 200 briefly, or by buffering the video signal for a period of time, for example. In the parental control system and corresponding method, a buffer large enough to contain 10 seconds of video signal is used; the multimedia memory 20 illustrated in Fig. 1 is a suitable memory.

In terms of segmentation, video input is segmented using cut detection, which is described in greater detail in, for example, in the aforementioned U.S. Patent No. 6,137,544. For every segment between two cuts, feature extraction is performed as described above. However in the obscure case when the visual segments are longer than n minutes (e.g.  $n > 2$  in the parental control system and corresponding method) then the system searches for audio clues to see if the granularity of audio segments is smaller. The smaller of the two is taken. The segments are then fed into the classification module. For each segment, a number representing the likelihood of the segment belonging to one of the sensitive categories is obtained; for multiple active filtering categories, multiple numbers are generated. It should be mentioned that the multimedia memory 20

advantageously can be sized to permit storage of many minutes of multimedia storage; it will be appreciated that adaptation of the parental control system to process segments which may be several minutes in length move the parental control system from the realm of real time processing and viewing into the realm of real time processing and near real time viewing.

5

Where material is deemed sensitive by the embodiment in multiple categories, a user configurable weighting system advantageously can be employed to bias the results of the filtering option. For example, if the user is configuring the parental control system to accommodate viewing by an older person, e.g., a grandparent, he or she may give more weight to the filtering category for violence (e.g., loud noises and blood make bring back traumatic memories), and less to the sexual content filter category (e.g., the grandparent has likely seen it all before, anyway). However, the user may employ a completely different weighting system for a teenager. When a user-set threshold for the total combined numerical rating is exceeded (including the calculations of all weighted material), the filtering module 290 is notified that it needs to filter the video signal.

10

15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995

However, to simplify the process for the filtering module 290, if the segmentation and categorization module 280 determines that consecutive segments  $S_n$  and  $S_{n+1}$  have a high likelihood of belonging to the same category, then these segments are preferably merged as a single output segment. It will be appreciated that the output from the segmentation process performed in module 280 is segments with a high likelihood of belonging to a particular sensitive category. Now, if consecutive segments belong to different categories, say "violence" followed by "profanity" then the union of the two segments can be marked for removal. If the system detects small gaps in between segments, say less than one minute, then the gap is also included in the segment for removal.

20

25

The filtering module 290, when it receives notification that it should act to remove a segment, makes a determination, based on the segment duration and preset configuration settings, of

the optimal method to be employed in filtering the offending content. For example, the module 290 can simply skip the segment, taking advantage of the aforementioned buffer. Alternatively, the filter module 290 advantageously can be configured to substitute another video signal for that segment (e.g., show a Barney the dinosaur interstitial or Web page). Moreover, assuming that the filter module 290 receives definitive information as to what part of a multimedia segment is to be removed, the filter module 290 advantageously can mask or blur out that particular portion. For example, when the audio portion of the segment contains an offensive word or phrase but is otherwise unobjectionable, the user may wish to merely garble the offending word or phrase rather than draw attention to the fact that a portion of the film was excised.

As mentioned above, the parental control system equipped or connected to a large multimedia memory would permit the controlling user to filter the original film in a myriad of ways to produce versions of the film suitable for several disparate groups in the user's household, since the system provides the capability of weighing or scaling the objectionable content so that the controlling user can decide which set of features are allowed under what conditions. For example, assume that there are two children in the house age 7 and 14 and the controlling user has identified different tolerance levels for the same content for each child. In that case, more filters would apply for the 7 year old and than would apply for the 14 year old. Thus, the parental control system can produce multilevel marking for the same movie so that the segments that are marked objectionable for the 14 year old are just a subset of those that are marked for the 7 year old.

It will be appreciated that although the learning phase has been completed and the classification phase has begun, the parental control system advantageously can receive feedback from the controlling user. That is, after the parental control system segments and classifies the movie, the user can be given a chance to review the marked segments and provide feedback to the system, identifying which segments are correctly marked or classified and which are marked

incorrectly. It will be appreciated that the next time that the system goes through the learning and classification phases, the system will produce better results.

From the discussion above, it will be appreciated that the parental control system according to the present invention provides the user with the capability of filtering a multimedia, e.g., a video, program based on user specified criteria in real time. Moreover, by merely increasing the size of a buffer memory coupled to other components of the parental control system, the user gains the ability to filter larger segments of the program in near real time. Advantageously, the parental control system according to the present invention provides structure by which a user can edit factory set criteria and/or enter additional criteria, for blocking or filtering of objectionable video program content.

It will also be appreciated that the parental control system according to the present invention advantageously provides a system which learns a user's preferences in terms of what types of content, or portions of that content, the user finds objectionable, and for what type of viewer such criteria are being provided, so that the system can then apply what it has learned in analyzing programs in the future. In an exemplary embodiment, the system is continually learning and fine tuning its behavior based on user input.

The parental control system described above provides circuitry by which features are extracted from the video program to assist the system in both teaching the learning component and in future filtering operations. It will be appreciated that "features" is a comprehensive term including, but not limited to, objects and their qualities from the video portion of the multimedia signal, sounds indicative of particular actions from the audio portion of the multimedia program, language from the audio portion of the multimedia program and/or from the close captioning data being sent along with the video portion of the multimedia program. In short, the parental control system includes a learning module and a filter module. The learning module advantageously can be

powered by one or more learning technologies, including Hidden Markov models and/or neural networks, into which criteria including factory set and user-supplied information is fed and combined with feature extraction data from a transcript feature extraction component, a video data feature extraction component, and an audio data feature extraction component. The learning module's resulting knowledge advantageously can be applied to a system filter module, which employs that knowledge to dynamically filter a video program's content according to the specified criteria.

Although presently preferred embodiments of the present invention have been described in detail hereinabove, it should be clearly understood that many variations and/or modifications of the basic inventive concepts herein taught, which may appear to those skilled in the pertinent art, will still fall within the spirit and scope of the present invention, as defined in the appended claims.